



Research Article

Computing Hitting Probabilities of Markov Chains: Structural Results with regard to the Solution Space of the Corresponding System of Equations

Hendrik Baumann ¹ and Thomas Hanschke ²

¹*Clausthal University of Technology, Institute of Mathematics, Erzstr. 1, 38678 Clausthal-Zellerfeld, Germany*

²Simulation Science Center Clausthal-Göttingen, Arnold-Sommerfeld-Str. 6, 38678 Clausthal-Zellerfeld, Germany

Correspondence should be addressed to Hendrik Baumann; hendrik.baumann@tu-clausthal.de

Received 18 October 2019; Accepted 6 December 2019; Published 13 January 2020

Academic Editor: Ying Hu

Copyright © 2020 Hendrik Baumann and Thomas Hanschke. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In a previous paper, we have shown that forward use of the steady-state difference equations arising from homogeneous discrete-state space Markov chains may be subject to inherent numerical instability. More precisely, we have proven that, under some appropriate assumptions on the transition probability matrix P , the solution space S of the difference equation may be partitioned into two subspaces $S = S_1 \oplus S_2$, where the stationary measure of P is an element of S_1 , and all solutions in S_1 are asymptotically dominated by the solutions corresponding to S_2 . In this paper, we discuss the analogous problem of computing hitting probabilities of Markov chains, which is affected by the same numerical phenomenon. In addition, we have to fulfill a somewhat complicated side condition which essentially differs from those conditions one is usually confronted with when solving initial and boundary value problems. To extract the desired solution, an efficient and numerically stable generalized-continued-fraction-based algorithm is developed.

1. Introduction

This paper is dedicated to homogeneous discrete-time Markov chains $X = (X_k)_{k \geq 0}$ with a countably infinite set of

states (labelled as $0, 1, 2, \dots$) and a one-step transition matrix $P = (P_{ij})_{0 \leq i, j < \infty}$ of the form

[illegible]

with some $n \geq 1$ and $P_{k,k+1} \neq 0$ for all $k \in \mathbb{N} = \{1, 2, 3, \dots\}$ and $P_{k,k-n} \neq 0$ for all $k \in \{n, n+1, n+2, \dots\}$. $P_{0\ell} = \delta_{0\ell}$ implies that 0 is an absorbing state, and the assumptions on $P_{k,k+1}$ and $P_{k,k-n}$ guarantee that state 0 is accessible from any state $k > 0$; that is, for all $k > 0$, there is some $m \in \mathbb{N}$ with $P_{k0}^{(m)} > 0$, where $(P_{ij}^{(m)}) = P^m$.

We shall calculate the sequence $(\alpha_k)_{k=0}^\infty$, where α_k denotes the probability that starting from $k > 0$ the Markov chain ultimately enters the absorbing state 0 (note that we only have assumed $\alpha_k > 0$ for all $k \in \mathbb{N}$. Although there is only one absorbing state, due to the infinite number of states, there is no guarantee that $\alpha_k = 1$); e.g., in a population model, this corresponds to eventual extinction of the population. Note that the choice $P_{0j} = \delta_{0j}$ is only due to simplicity. For arbitrary P_{0j} , state 0 is not absorbing anymore, but these values do not have any impact on the probabilities α_k . Since they still reflect the probabilities that the state 0 is eventually reached, they are referred to as *hitting probabilities*.

The sequence $(\alpha_k)_{k=0}^\infty$ is the uniquely determined minimal nonnegative solution to

$$\begin{aligned} x_k &= \sum_{\ell=0}^{\infty} P_{k\ell} x_\ell, \quad k = 1, 2, 3, \dots, \\ x_0 &= 1, \end{aligned} \quad (2)$$

for example, see Section 8.4 in [1]. Due to the structure of P given in (1), we conclude that $(\alpha_k)_{k=0}^\infty$ is the minimal nonnegative solution to the $(n+1)$ th-order homogeneous linear recurrence relation

$$\begin{aligned} P_{k+n,k+n+1} x_{k+n+1} + (P_{k+n,k+n} - 1) x_{k+n} + \dots + P_{k+n,k} x_k &= 0, \\ k &= 0, 1, 2, \dots, \end{aligned} \quad (3)$$

subject to the side conditions

$$\begin{cases} x_0 = 1, \\ x_1 = P_{10} x_0 + P_{11} x_1 + P_{12} x_2, \\ x_2 = P_{20} x_0 + P_{21} x_1 + P_{22} x_2 + P_{23} x_3, \\ \vdots \\ x_{n-1} = P_{n-1,0} x_0 + \dots + P_{n-1,n-1} x_{n-1} + P_{n-1,n} x_n. \end{cases} \quad (4)$$

Since we assume $P_{k+n+1,k+n} \neq 0$ for all k , system (4) has rank n , such that the desired solution $(\alpha_k)_{k=0}^\infty$ is uniquely determined by α_1 ((4) and (3)).

The solution of the difference scheme involves two problems:

- (i) To start the sequence $(\alpha_k)_{k=0}^\infty$, we need α_1 . Put $\alpha_k = x_k^{(1)} \alpha_1 + x_k^{(2)}$ for $k = 0, 1, 2, \dots$, where $x_0^{(1)} = 0$, $x_1^{(1)} = 1$, $x_0^{(2)} = 1$, and $x_1^{(2)} = 0$ (we will refer to this representation of $(\alpha_k)_{k=0}^\infty$ as (FW1)). Due to the linear structure of (3) and (4), $(x_k^{(1)})_{k=0}^\infty$ and $(x_k^{(2)})_{k=0}^\infty$ satisfy (3) and (4) up to the first two equations. Another representation of $(\alpha_k)_{k=0}^\infty$ is $\alpha_k = \tilde{x}_k^{(1)} c + \tilde{x}_k^{(2)} (1 - c)$ with $x_k^{(1)} = 1$ for $k = 0, 1, 2, \dots$, and another solution

$(x_k^{(2)})_{k=0}^\infty$ to (3) and (4) being linearly independent of $(x_k^{(1)})_{k=0}^\infty$. This representation (which we will refer to as (FW2)) results from the fact that P is stochastic.

If the difference equation can be solved analytically, α_1 and c , respectively, can be determined by the requirement that $(\alpha_k)_{k=0}^\infty$ becomes nonnegative and minimal. To prove this assertion, knowledge of all solutions to (3) and (4) (including their asymptotic behaviour) is necessary. But if the determination of $(\alpha_k)_{k=0}^\infty$ is based on numerical calculations such that only a finite number of components of $(\alpha_k)_{k=0}^\infty$ can be checked, these proofs fail. In this situation, the determination of α_1 and c , respectively, is an open problem.

- (ii) From the theory of linear difference equations (see Miller [2]), it is known that there exists $n+1$ linearly independent functions $x^{(0)}, x^{(1)}, \dots, x^{(n)}$ defined on \mathbb{N}_0 which assume prescribed values at $k = 0, 1, \dots, n$ and satisfy the homogeneous recurrence relation (3) for all $k \geq 0$. In the sequel, it is shown that the solution space S of the recurrence relation (3) is the direct sum

$$S = S_1 \oplus S_2, \quad (5)$$

of two subspaces S_1 and S_2 with the property that every solution $y \in S_2$ dominates over every solution $x \in S_1$, i.e.,

$$\lim_{k \rightarrow \infty} \frac{x_k}{y_k} = 0. \quad (6)$$

In addition, we have $\dim S_1 = n$, $\dim S_2 = 1$, and $\alpha = (\alpha_k)_{k=0}^\infty \in S_1$. From the literature [3–7], it is known that computing solutions $x \in S_1$ by means of forward computation is not a meaningful procedure since forward computation of a subdominant solution becomes numerically unstable. The reason is as follows: if we replace $x_{k_0}, \dots, x_{k_0+n-1}$ by approximate values $z_{k_0}, \dots, z_{k_0+n-1}$ (due to rounding errors, for example), we would have $z_k = \alpha x_k + \beta y_k$ for all $k \geq k_0$ with some $\beta > 0$ and $y \in S_2$. Even with infinite precision, this would imply $|(z_k - x_k)/x_k| \rightarrow \infty$ as $k \rightarrow \infty$; that is, the relative error becomes arbitrarily large (see [4] for more details in the case $n = 2$). Since rounding errors are inevitable, forward computation of $x \in S_1$ will always fail.

In [8, 9], it was shown that the problem of numerical instability also arises in the context of computing invariant measures of time-homogeneous and discrete-state space Markov chains with band-structured transition probability matrices. The idea was to construct a decoupled recursion for the solutions $x \in S_1$ in which the dominating solutions $y \in S_2$ do not appear. To carry out the reduction of order for the linear difference equation (3), we apply a technique which relies on generalized-continued fractions. Since the resulting linear difference equation is of order n , instead of $n+1$, the problem of finding the initial value α_1 (or the value c) becomes obsolete.

2. Generalized-Continued Fractions and Linear Difference Equations

For the following, we need some results on generalized-continued fractions (GCFs). To point out the analogy to our results in [8], we use the same notations and terminologies.

Definition 1. A generalized-continued fraction (GCF) of dimension n is an $(n+1)$ -tuple $(a^{(1)}, \dots, a^{(n)}, b)$ of real-valued sequences and a convergence structure as follows. Let $A^{(1)}, \dots, A^{(n)}, B$ be the principal solution of the corresponding homogeneous linear difference equation:

$$x_{k+n+1} = b_k x_{k+n} + a_k^{(n)} x_{k+n-1} + \dots + a_k^{(1)} x_k, \quad k = 0, 1, 2, \dots, \quad (7)$$

satisfying

$$\begin{cases} A_j^{(i)} = \delta_{i,j+1} = \begin{cases} 1 & \text{for } i = j+1 \\ 0 & \text{for } i \neq j+1 \end{cases}, \quad i = 1, \dots, n; j = 0, \dots, n, \\ B_j = 0, \quad \text{for } j = 0, \dots, n-1, \\ B_n = 1. \end{cases} \quad (8)$$

The GCF is said to converge iff all the limits

$$\xi^{(i)} = \lim_{k \rightarrow \infty} \frac{A_k^{(i)}}{B_k}, \quad i = 1, \dots, n, \quad (9)$$

do exist.

Convergence of a GCF is indicated by the notation

$$\begin{bmatrix} \xi^{(1)} \\ \vdots \\ \xi^{(n)} \end{bmatrix} \stackrel{k}{=} \begin{bmatrix} a_0^{(1)} & a_1^{(1)} & a_2^{(1)} & \dots \\ \dots & \dots & \dots & \dots \\ a_0^{(n)} & a_1^{(n)} & a_2^{(n)} & \dots \\ b_0 & b_1 & b_2 & \dots \end{bmatrix}. \quad (10)$$

Terminating a GCF after the N -th column, we get the so-called N -th approximants:

$$\begin{bmatrix} \frac{A_{N+n+1}^{(1)}}{B_{N+n+1}} \\ \vdots \\ \frac{A_{N+n+1}^{(n)}}{B_{N+n+1}} \end{bmatrix} = \begin{bmatrix} a_0^{(1)} & a_1^{(1)} & \dots & a_N^{(1)} \\ \dots & \dots & \dots & \dots \\ a_0^{(n)} & a_1^{(n)} & \dots & a_N^{(n)} \\ b_0 & b_1 & \dots & b_N \end{bmatrix}. \quad (11)$$

The term “generalized-continued fraction” becomes obvious if forward computation of the N -th approximant of a GCF is replaced by the equivalent backward algorithm. It is known that the recurrence relation (7) can be converted into a first-order vector recursion of the form

$$u_{k+1} = W_k u_k, \quad k = 0, 1, 2, \dots, \quad (12)$$

by putting

$$u_k = \begin{bmatrix} x_k \\ x_{k+1} \\ \vdots \\ x_{k+n-1} \\ x_{k+n} \end{bmatrix}, \quad W_k = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & 1 \\ a_k^{(1)} & a_k^{(2)} & \dots & \dots & a_k^{(n)} & b_k \end{bmatrix}, \quad (13)$$

$$k = 0, 1, 2, \dots$$

Comparing (7) and (12), it is seen that the numerators $A_{N+n+1}^{(1)}, \dots, A_{N+n+1}^{(n)}$ and the denominators B_{N+n+1} of the GCF appear in the last row of the matrix:

$$W_N W_{N-1} \dots W_0. \quad (14)$$

Consider now the backward recurrence scheme

$${}_{(N)}v_k = {}_{(N)}v_{k+1} W_k, \quad k = N, N-1, \dots, 0, \quad (15)$$

with initial vector ${}_{(N)}v_{N+1} = (0, \dots, 0, 1)$. Then,

$${}_{(N)}v_0 = {}_{(N)}v_{N+1} W_N W_{N-1} \dots W_0 = (A_{N+n+1}^{(1)}, \dots, A_{N+n+1}^{(n)}, B_{N+n+1}). \quad (16)$$

Writing (15) in the expanded form and inserting the structure of W_k , we get

$$\begin{aligned} {}_{(N)}v_k^{(1)} &= {}_{(N)}v_{k+1}^{(n+1)} a_k^{(1)}, \quad k = N, N-1, \dots, 0, \\ {}_{(N)}v_k^{(i)} &= {}_{(N)}v_{k+1}^{(i-1)} + {}_{(N)}v_{k+1}^{(n+1)} a_k^{(i)}, \quad i = 2, \dots, n, k = N, N-1, \dots, 0, \\ {}_{(N)}v_k^{(n+1)} &= {}_{(N)}v_{k+1}^{(n)} + {}_{(N)}v_{k+1}^{(n+1)} b_k, \quad k = N, N-1, \dots, 0, \end{aligned} \quad (17)$$

or equivalently

$$\begin{cases} (N)r_k^{(1)} := \frac{(N)v_k^{(1)}}{(N)v_k^{(n+1)}} = \frac{a_k^{(1)}}{b_k + (N)r_{k+1}^{(n)}}, & k = N, N-1, \dots, 0, \\ (N)r_k^{(i)} := \frac{(N)v_k^{(i)}}{(N)v_k^{(n+1)}} = \frac{a_k^{(i)} + (N)r_{k+1}^{(i-1)}}{b_k + (N)r_{k+1}^{(n)}}, & i = 2, \dots, n; k = N, N-1, \dots, 0, \end{cases} \quad (18)$$

with initial values $(N)r_{N+1}^{(i)} = 0$ for $i = 1, \dots, n$. Hence,

$$\frac{A_{N+n+1}^{(i)}}{B_{N+n+1}} = (N)r_0^{(i)}, \quad i = 1, \dots, n, \quad (19)$$

$$\xi^{(i)} = \lim_{N \rightarrow \infty} (N)r_0^{(i)}, \quad i = 1, \dots, n.$$

Alternative procedures for computing GCFs are described in [10].

The relations between GCFs and linear difference equations have been first recognized by Perron [11]. Perron's results were generalized by Van der Cruyssen [10], Hanschke [12, 13], and Levrie and Bultheel [14]. The following theorem is due to Van der Cruyssen [10].

Theorem 1. A GCF $(a^{(1)}, \dots, a^{(n)}, b)$ converges iff there are $n+1$ linearly independent solutions $x^{(1)}, \dots, x^{(n)}, y$ of the recurrence relation (7) satisfying

$$\lim_{k \rightarrow \infty} \frac{x_k^{(i)}}{y_k} = 0, \quad i = 1, \dots, n, \quad (20)$$

$$\begin{vmatrix} x_0^{(1)} & \dots & x_{n-1}^{(1)} \\ \vdots & & \\ x_0^{(n)} & \dots & x_{n-1}^{(n)} \end{vmatrix} \neq 0.$$

As pointed out by Gautschi [4, 5] and Van der Cruyssen [10], forward computation of a solution $x \in \text{span}(x^{(1)}, \dots, x^{(n)})$ is numerically unstable. Van der Cruyssen [10] establishes that if the limits

$$\begin{bmatrix} \xi_l^{(1)} \\ \vdots \\ \xi_l^{(n)} \end{bmatrix} = \begin{bmatrix} a_l^{(1)} & a_{l+1}^{(1)} & a_{l+2}^{(1)} & \dots \\ \dots & \dots & \dots & \dots \\ a_l^{(n)} & a_{l+1}^{(n)} & a_{l+2}^{(n)} & \dots \\ b_l & b_{l+1} & b_{l+2} & \dots \end{bmatrix}, \quad (21)$$

exist for all $l \geq 0$, then $x \in \text{span}(x^{(1)}, \dots, x^{(n)})$ iff

$$x_{l+n} = -\xi_l^{(n)} x_{l+n-1} - \dots - \xi_l^{(1)} x_l, \quad l = 0, 1, 2, \dots \quad (22)$$

Combining (18), (19), and (22), one obtains an efficient algorithm (which we will refer to as *Van der Cruyssen's algorithm*) for approximating the first $L+1$ components x_0, \dots, x_L of an element $x \in \text{span}(x^{(1)}, \dots, x^{(n)})$ with prescribed values x_0, \dots, x_{n-1} :

Step 1. Select $N > L - n$ and define $(N)r_l^{(i)}$ for $i = 1, \dots, n; l = 0, 1, \dots, N+1$ by

$$(N)r_{N+1}^{(1)} = \dots = (N)r_{N+1}^{(n)} = 0,$$

$$(N)r_l^{(1)} = \frac{a_l^{(1)}}{b_l + (N)r_{l+1}^{(n)}}, \quad l = N, N-1, \dots, 0,$$

$$(N)r_l^{(i)} = \frac{a_l^{(i)} + (N)r_{l+1}^{(i-1)}}{b_l + (N)r_{l+1}^{(n)}}, \quad i = 2, \dots, n; l = N, N-1, \dots, 0. \quad (23)$$

Step 2. Set

$$\begin{aligned} (N)x_0 &= x_0, \dots, (N)x_{n-1} = x_{n-1}, \\ (N)x_{l+n} &= -(N)r_l^{(n)} (N)x_{l+n-1} - \dots - (N)r_l^{(1)} (N)x_l, \end{aligned} \quad (24)$$

for $l = 0, 1, \dots, L - n$.

Step 3. Increase N until the accuracy of the iterates is within prescribed limits.

For any l , the vector $((N)r_l^{(1)}, \dots, (N)r_l^{(n)})^T$ is an approximant of a GCF. Hence, convergence of the algorithm is related to convergence of GCFs. For the latter, we cite a result which can be interpreted as a generalization of Pringsheim's convergence criterion ([15], p. 58) for ordinary continued fractions.

Theorem 2 (Levrie [16] and Perron [17]). *The GCF*

$$\begin{bmatrix} a_0^{(1)} & a_1^{(1)} & a_2^{(1)} & \dots \\ \dots & \dots & \dots & \dots \\ a_0^{(n)} & a_1^{(n)} & a_2^{(n)} & \dots \\ b_0 & b_1 & b_2 & \dots \end{bmatrix}, \quad (25)$$

converges if it satisfies the so-called dominance condition, i.e.,

$$|a_k^{(1)}| + |a_k^{(2)}| + \dots + |a_k^{(n)}| + 1 \leq |b_k|, \quad \text{for all } k \geq 0. \quad (26)$$

3. Main Results

To make (3) congruent with (7), we put

$$a_k^{(i)} = \frac{P_{k+n,k+i-1}}{P_{k+n,k+n+1}}, \quad i = 1, \dots, n, \quad k = 0, 1, 2, \dots, \quad (27)$$

$$b_k = -\frac{P_{k+n,k+n} - 1}{P_{k+n,k+n+1}}, \quad k = 0, 1, 2, \dots \quad (28)$$

Theorem 3. Let $a_k^{(i)}, b_k$ be defined as in (27) and (28), and let $P_{k,k+1} \neq 0$ for all $k \in \mathbb{N}$ and $P_{k,k-n} \neq 0$ for all $k \geq n$. Then, the limits

$$\begin{bmatrix} \xi_l^{(1)} \\ \vdots \\ \xi_l^{(n)} \end{bmatrix} = \begin{bmatrix} a_l^{(1)} & a_{l+1}^{(1)} & a_{l+2}^{(1)} & \dots \\ \dots & \dots & \dots & \dots \\ a_l^{(n)} & a_{l+1}^{(n)} & a_{l+2}^{(n)} & \dots \\ b_l & b_{l+1} & b_{l+2} & \dots \end{bmatrix}, \quad (29)$$

exist for all $l \geq 0$. The vector $(\alpha_k)_{k=0}^\infty$ of the absorption probabilities with respect to state 0 is a dominated solution of (7) and satisfies the decoupled recursion

$$x_{l+n} = -\xi_l^{(n)} x_{l+n-1} - \dots - \xi_l^{(1)} x_l, \quad l = 0, 1, 2, \dots \quad (30)$$

Proof. As in the corresponding statement Theorem 3 in [8], the basic idea for the proof is that $(\alpha_k)_{k=0}^\infty$ is the limit of the solutions of appropriately truncated systems of equations. We give some details on the probabilistic interpretation of these systems although these considerations are well-known from standard literature on Markov chains (e.g., [18]).

By ${}^{(N)}P = (P_{ik})_{0 \leq i, k \leq N}$, denote the northwest corner truncation of P , and by $X^{(N)} = (x_k^{(N)})_{k \geq 0}$, denote a discrete-time Markov chain with state space $\{0, 1, \dots, N+1\}$ and block-transition probability matrix

$$\begin{pmatrix} {}^{(N)}P & q \\ 0 & 1 \end{pmatrix}, \quad (31)$$

where $q = (q_i)_{i=0}^N$ is a column vector with $q_i = \sum_{k=N+1}^\infty P_{ik}$. For $X^{(N)}$, both 0 and $N+1$ are absorbing states. By $\alpha_k^{(N)}$, denote the probability that $X^{(N)}$ will be absorbed in state 0 when starting in state k . Obviously, we have $\alpha_{N+1}^{(N)} = 0$, and hence, the values $\alpha_k^{(N)}$ define a solution to the truncated system:

$$\begin{aligned} {}^{(N)}x_0 &= 1, \\ {}^{(N)}x_k &= \sum_{\ell=0}^N P_{k\ell} \cdot {}^{(N)}x_\ell, \\ k &= 1, \dots, N. \end{aligned} \quad (32)$$

Since for the original Markov chain X , state 0 was assumed to be accessible from all $k > 0$, either state 0 or $N+1$ is accessible from all states $k \in \{1, \dots, N\}$ for $X^{(N)}$. Hence, every state $\in \{1, \dots, N\}$ is transient, implying that ${}^{(N)}I - (P_{k\ell})_{k, \ell=1}^N$ is invertible, where ${}^{(N)}I$ is the $N \times N$ -identity matrix. Thus, $(\alpha_k^{(N)})_{k=0}^N$ is the unique solution to (32).

Under the given assumptions on the structure of P , (32) coincides with the homogeneous linear difference equation (3) (or (7) in terms of b_k and $a_k^{(i)}$), coupled to the original

side conditions (4) (assuming that $N > n$) and the additional side condition:

$${}^{(N)}x_{N+1} = 0. \quad (33)$$

For the GCF $\begin{bmatrix} \xi_l^{(1)} \\ \vdots \\ \xi_l^{(n)} \end{bmatrix}$ defined in (29), let us denote the numerators by $A_{l,k}^{(1)}, \dots, A_{l,k}^{(n)}$ and the denominators by $B_{l,k}$. Then, $(A_{l,k}^{(i)})_k$ and $(B_{l,k})_k$ satisfy

$$x_{k+n+1} = b_{k+l} x_{k+n} + a_{k+l}^{(n)} x_{k+n-1} + \dots + a_{k+l}^{(1)} x_k, \quad k = 0, 1, 2, \dots, \quad (34)$$

and build up a fundamental system of solutions to equation (7) for $k \geq l$. Hence, any solution of the recurrence relation (7) can be expressed in terms of these functions. In view of their initial values

$$\begin{cases} A_{l,j}^{(i)} = \delta_{i,j+1} = \begin{cases} 1, & i = j+1 \\ 0, & i \neq j+1 \end{cases}, & i = 1, \dots, n; \quad j = 0, \dots, n, \\ B_{l,j} = 0, & \text{for } j = 0, \dots, n-1, \\ B_{l,l+n} = 1, \end{cases} \quad (35)$$

we get

$${}^{(N)}x_k = {}^{(N)}x_l A_{l,k-l}^{(1)} + \dots + {}^{(N)}x_{l+n-1} A_{l,k-l}^{(n)} + {}^{(N)}x_{l+n} B_{l,k-l}, \quad (36)$$

$$k = l, l+1, \dots$$

Choosing $k = N+1$ and utilizing (33), equation (36) reduces to

$${}^{(N)}x_{l+n} B_{l,N-l+1} = -{}^{(N)}x_l A_{l,N-l+1}^{(1)} - \dots - {}^{(N)}x_{l+n-1} A_{l,N-l+1}^{(n)}. \quad (37)$$

Dividing both sides of (37) by $B_{l,N-l+1}$ and ${}^{(N)}x_0 = 1$, we formally arrive at

$${}^{(N)}x_{l+n} = -{}^{(N)}x_l \frac{A_{l,N-l+1}^{(1)}}{B_{l,N-l+1}} - \dots - {}^{(N)}x_{l+n-1} \frac{A_{l,N-l+1}^{(n)}}{B_{l,N-l+1}}. \quad (38)$$

Passing to the limit $N \rightarrow \infty$, we get the decoupled recursion (30), provided that the limits $\lim_{N \rightarrow \infty} (A_{l,N-l+1}^{(i)} / B_{l,N-l+1})$ do exist for $i = 0, 1, \dots, n$, and that $\lim_{N \rightarrow \infty} {}^{(N)}x_k = \alpha_k$ for all $k \in \mathbb{N}_0$.

(i) Using the definitions of $a_k^{(i)}$ and b_k and the fact that P is stochastic, we immediately find

$$\begin{aligned} |a_k^{(1)}| + |a_k^{(2)}| + \dots + |a_k^{(n)}| + 1 &= \frac{\sum_{i=1}^n P_{k+n,k+i-1} + P_{k+n,k+n+1}}{P_{k+n,k+n+1}} \\ &= \frac{1 - P_{k+n,k+n}}{P_{k+n,k+n+1}} = |b_k|. \end{aligned} \quad (39)$$

- (ii) Hence, Theorem 2 guarantees that the limits $\lim_{N \rightarrow \infty} (A_{l,N-l+1}^{(i)} / B_{l,N-l+1})$ do exist.
- (iii) As pointed out above, the unique solution to (3) and (4) with the additional side condition (33) is given by ${}^{(N)}x_k = \alpha_k^{(N)}$, where for the truncated Markov chain $X^{(N)}$, $\alpha_k^{(N)}$ is the probability to enter state 0 before $N+1$ when starting in k . For the original Markov chain X , let τ_0 and τ_{N+1} be the first hitting times of 0 and $N+1$, respectively, that is

$$\tau_i = \inf\{m \in \mathbb{N}: X_m = i\}. \quad (40)$$

- (iv) Then, we can write $\alpha_k^{(N)} = \mathbb{P}(\tau_0 < \tau_{N+1} | X_0 = k)$. For $N \rightarrow \infty$, these probabilities increase monotonically up to the limit $\mathbb{P}(\tau_0 < \infty | X_0 = k) = \alpha_k$. Hence,

$$\lim_{N \rightarrow \infty} {}^{(N)}x_k = \lim_{N \rightarrow \infty} \alpha_k^{(N)} = \alpha_k. \quad (41)$$

□

Remark 1. Notice that the statement of Theorem 3 consists of three parts:

- (i) The sequence of absorption probabilities is dominated by other solutions of the difference equation (3).

- (ii) By means of GCFs, we are able to construct a decoupled recursion which is not affected by numerical instability.
- (iii) Since the decoupled recursion is a linear difference equation of order n (instead of order $n+1$), the problem of finding the initial value α_1 (or the value c) is eliminated since $\alpha_1, \dots, \alpha_n$ can be obtained by exploiting (30) for $l=0$ and the side conditions (4). In other words, the calculation of α_1 or c is incorporated in Van der Cruyssen's algorithm.

The fact that the solution of the truncated system converge to the desired minimal solution is far from new. But to the best of our knowledge, although the numerical solution of the system is associated with massive problems, both the derivation of the structural result $S_1 \oplus S_2$ and the relationship to generalized-continued fractions have not been addressed in the literature before (up to the corresponding considerations in [8] concerning the computation of invariant measures of Markov chains).

4. Continuous-Time Markov Chains

The results of the preceding section can easily be extended to continuous-time Markov chains $Y = (Y_t)_{t \geq 0}$ generated by a conservative, irreducible, and regular Q -matrix of the form

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & \dots \\ Q_{10} & Q_{11} & Q_{12} & 0 & \dots & \\ Q_{20} & Q_{21} & Q_{22} & Q_{23} & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\ Q_{n0} & Q_{n1} & \dots & Q_{n,n-1} & Q_{nm} & Q_{n,n+1} & 0 & \dots \\ 0 & Q_{n+1,1} & Q_{n+1,2} & \dots & Q_{n+1,n} & Q_{n+1,n+1} & Q_{n+1,n+2} & 0 & \dots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}, \quad (42)$$

where $n \geq 1$, $Q_{kk} \neq 0$, and $Q_{k,k+1} \neq 0$ for all $k \in \{1, 2, 3, \dots\}$ and $Q_{k,k-n} \neq 0$ for all $k \in \{n, n+1, n+2, \dots\}$.

For the continuous-time Markov chain Y , the probability α_k of eventually reaching state 0 when starting in state k is equal to the probability of eventually reaching state 0 for the embedded jump chain $X = (X_m)_{m=0}^{\infty}$ with transition probability matrix $P = (P_{k\ell})_{k,\ell=0}^{\infty}$ where

$$P_{k\ell} = \begin{cases} \delta_{k\ell}, & Q_{k\ell} = 0 (\iff k = 0), \\ \frac{Q_{k\ell}}{-Q_{kk}} + \delta_{k\ell}, & Q_{k\ell} \neq 0. \end{cases} \quad (43)$$

Hence, α_k solves the homogeneous system (3) with the side conditions (4), where $P_{k\ell}$ is defined as above; that is, we obtain the modified homogeneous linear difference equation:

$$Q_{k+n,k+n+1}x_{k+n+1} + Q_{k+n,k+n}x_{k+n} + \cdots + Q_{k+n,k}x_k = 0, \quad k = 0, 1, 2, \dots, \quad (44)$$

subject to the side conditions

$$\begin{cases} x_0 = 1, \\ 0 = Q_{10}x_0 + Q_{11}x_1 + Q_{12}x_2, \\ 0 = Q_{20}x_0 + Q_{21}x_1 + Q_{22}x_2 + Q_{23}x_3, \\ \vdots \\ 0 = Q_{n-1,0}x_0 + \cdots + Q_{n-1,n-1}x_{n-1} + Q_{n-1,n}x_n. \end{cases} \quad (45)$$

In view of the difference equation (3) on which the definition of GCFs relies, we set

$$\begin{aligned} a_k^{(i)} &= \frac{Q_{k+n,k+i-1}}{Q_{k+n,k+n+1}}, \quad i = 1, \dots, n, \\ b_k &= \frac{Q_{k+n,k+n}}{Q_{k+n,k+n+1}}. \end{aligned} \quad (46)$$

Using (43), we obtain that this corresponds exactly to the choice of $a_k^{(i)}$ and b_k given by (27) and (28), respectively. Therefore, Theorem 3 can be directly transferred to continuous-time Markov chains.

Theorem 4. Let $a_k^{(i)}, b_k$ be defined as in (46) and let $Q_{k,k+1} \neq 0$ for all $k \in \mathbb{N}$ and $Q_{k,k-n} \neq 0$ for all $k \geq n$. Then, the limits

$$\begin{bmatrix} \xi_l^{(1)} \\ \vdots \\ \xi_l^{(n)} \end{bmatrix} = \begin{bmatrix} a_l^{(1)} & a_{l+1}^{(1)} & a_{l+2}^{(1)} & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_l^{(n)} & a_{l+1}^{(n)} & a_{l+2}^{(n)} & \cdots \\ b_l & b_{l+1} & b_{l+2} & \cdots \end{bmatrix}, \quad (47)$$

exist for all $l \geq 0$. The vector $(\alpha_k)_{k=0}^\infty$ of absorption probabilities with respect to state 0 is a dominated solution of (7) and satisfies the decoupled recursion:

$$x_{l+n} = -\xi_l^{(n)} x_{l+n-1} - \cdots - \xi_l^{(1)} x_l, \quad l = 0, 1, 2, \dots \quad (48)$$

5. Numerical Examples

As a simple example, we consider a modification of the famous gambler's ruin problem: A gambler gambles against a bank (with infinite capital) and wins or loses a coin in each round with probability p and $1-p$, respectively. Let X_k be the number of coins the gambler owns after n rounds. Then, $(X_k)_{k \geq 0}$ is a discrete-time Markov chain with the above transition structure where $n = 1$, $P_{i,i-1} = 1-p$, $P_{ii} = 0$, and $P_{i,i+1} = p$ for all $i \geq 1$. Due to $n = 1$, (4) simplifies to $\alpha_0 = 1$, and (3) reduces to

$$px_{k+2} - x_{k+1} + (1-p)x_k = 0, \quad k = 0, 1, 2, \dots \quad (49)$$

We briefly demonstrate how to determine the exact solution: As pointed out above (representation (FW2)), a solution is given by the constant $\tilde{x}_k^{(1)} = 1$ for $k = 0, 1, 2, \dots$

For $p = (1/2)$, another linearly independent solution is given by $\tilde{x}_k^{(2)} = k$, and hence, $\alpha_k = c + (1-c)k$ for $k = 0, 1, 2, \dots$. Due to nonnegativity and minimality, we have $c = 1$, and hence, $\alpha_k = 1$ for all $k = 0, 1, 2, \dots$. For $p < (1/2)$, another solution is given by $\tilde{x}_k^{(2)} = ((1-p)/p)^k$, and hence, $\alpha_k = c + (1-c)((1-p)/p)^k$ for $k = 0, 1, 2, \dots$. Due to $((1-p)/p) > 1$, minimality and nonnegativity imply $c = 1$, and hence, still $\alpha_k = 1$ for $k = 0, 1, 2, \dots$. For $p > (1/2)$, we still have $\alpha_k = c + (1-c)((1-p)/p)^k$, but here, due to $((1-p)/p) < 1$, we have to choose $c = 0$ and obtain $\alpha_k = ((1-p)/p)^k$ for $k = 0, 1, 2, \dots$. Regarding the analytic derivation of the exact solution, we could use the representation (FW1) in a similar way.

In most applications, it is not possible to find the explicit representation of α_k , and therefore, we have to use numerical calculations. For the sake of simplicity and in order to be able to compare our results with the exact solution, we apply numerical computation first to the gamblers's-ruin example. Using data type double in C++, for $p = (1/3)$, for the values $(x_k^{(1)})_{k=0}^\infty$ and $(x_k^{(2)})_{k=0}^\infty$ in the representation (FW1), we obtain the results depicted in Table 1.

Note that the knowledge that $\alpha_k = x_k^{(1)} + x_k^{(2)}$ relies on $\alpha_1 = 1$. So even if we knew the exact value for α_1 (which we usually do not know if we use numerical computations), we can get extremely wrong results from applying forward computation of the representation (FW1). In the representation (FW2), we can use $\tilde{x}_k^{(2)} = x_k^{(1)}$ or $\tilde{x}_k^{(2)} = x_k^{(2)}$. For both choices, the numerically computed values suggest that $\tilde{x}_k^{(2)}$ increases exponentially, and hence, for obtaining the minimal nonnegative solution, we should choose $c = 1$ and $c = 0$. So here, using numerically computed values for FW2 works. However, it is obvious that this argument is susceptible to misinterpretations.

The GCF-based approach requires a truncation level N . For $N = 120$, we obtain $\alpha_k = 1$ for $k \leq 100$ up to errors $< 10^{-6}$.

We demonstrate the use of the GCF-based approach in more detail by applying it in a situation where we do not know the exact solution. We will see that the method provides plausible results. Let $n = 2$ and

$$\begin{aligned} P_{k,k+1} &= \frac{1}{2} + g(k), \quad k = 1, 2, \dots, \\ P_{kk} &= \frac{1}{4}, \quad k = 1, 2, \dots, \\ P_{10} &= \frac{1}{4} - g(k), \\ P_{k,k-1} &= \frac{1}{8} - g(k), \quad k = 2, 3, \dots, \\ P_{k,k-2} &= \frac{1}{8}, \quad k = 2, 3, \dots \end{aligned} \quad (50)$$

For $g(k) = 0$, $g(k) = (1/2)^{k+2}$, and $g(k) = (1/8k)$, some values can be found in Table 2. For $N = 120$ and $N = 150$, the first six digits of the values $\alpha_0, \dots, \alpha_{100}$ coincide which

TABLE 1: Forward computation applied to the gambler's ruin problem.

k	$x_k^{(1)}$	$x_k^{(2)}$	$\alpha_k = x_k^{(1)} + x_k^{(2)}$
0	0	1	1
1	1	0	1
2	3	-2	1
3	7	-6	1
4	15	-14	1
10	1023	-1022	1
20	1048575	-1048574	1
30	1073741823	-1073741822	1
40	1099511627774.999878	-1099511627774.000244	0.999634
50	1125899906842622.875	-1125899906842622.250	0.625
75	.	.	-25165824
100	.	.	-844424930131968

TABLE 2: Numerical computation of the absorption probabilities by means of the GCF-based approach.

k	α_k for $g(k) = 0$ (exact)	α_k for $g(k) = 0$	α_k for $g(k) = (1/2^{k+2})$	α_k for $g(k) = (1/8k)$
0	1.000000	1.000000	1.000000	1.000000
1	0.750000	0.750000	0.579198	0.545488
2	0.625000	0.625000	0.495038	0.454585
3	0.500000	0.500000	0.373473	0.323282
4	0.406250	0.406250	0.303615	0.251803
5	0.328125	0.328125	0.242391	0.191475
6	0.265625	0.265625	0.195995	0.148602
8	0.173828	0.173828	0.127894	0.089963
10	0.113770	0.113770	0.083650	0.055233
15	0.039429	0.039429	0.028984	0.016904
20	0.013665	0.013665	0.010045	0.005333
50	0.000024	0.000024	0.000017	0.000007

shows that our results are trustworthy. Furthermore, as pointed out above, we can check the results with respect to plausibility: For $g(k) = 0$, we can compute the exact solution by means of the theory of linear difference equations with constant coefficients and obtain

$$\alpha_k = \frac{2 + \sqrt{5}}{2\sqrt{5}} \left(\frac{1 + \sqrt{5}}{4} \right)^k - \frac{2 - \sqrt{5}}{2\sqrt{5}} \left(\frac{1 - \sqrt{5}}{4} \right)^k, \quad k = 0, 1, 2, \dots, \quad (51)$$

see first column of Table 2. In general, the term $g(k)$ may be interpreted as a kind of disturbance. For $g(k) > 0$, the effect of the Markov chain “drifting away” from state 0 is enforced. In particular, since we have $\lim_{k \rightarrow \infty} g(k) = 0$ for our choices of g , this effect is strong for small states k . Hence, for small k , α_k decreases quite slowly for $g(k) = 0$ and more quickly for the other choices. Since $(1/8k) > (1/2^{k+2})$, we observe that α_k decreases faster for $g(k) = (1/8k)$.

6. Conclusion and Further Research

This paper deals with the evaluation of higher-order linear difference equations associated with the computation of hitting probabilities of discrete-state space Markov chains. To extinguish effects of inherent instability, a generalized-continued-fraction algorithm is established. Although aspects of inherent instability have been widely discussed in the

literature (e.g., [3–7]), their relevance to Markov chains has not been considered yet. It is to be expected that much more problems in stochastic modeling are subject to this phenomenon.

Similar to the computation of absorption or hitting probabilities, mean hitting times can be calculated: let $\alpha_i = 1$ and $X_0 = i$. Then, one may inquire for the mean time m_i which elapses until state 0 is assumed. With $m_0 = 0$, we have

$$\begin{aligned} m_i &= 1 + \sum_{j=0}^{\infty} p_{ij} m_j, \\ \text{or } 0 &= 1 + \sum_{j=0}^{\infty} q_{ij} m_j. \end{aligned} \quad (52)$$

Obviously, system (52) for $(m_i)_{i=0}^{\infty}$ is an inhomogeneous version of the homogeneous system for $(\alpha_i)_{i=0}^{\infty}$. If the transition structure is given by (1) or (42), respectively, we obtain an inhomogeneous version of the difference equation (3), which is subject to numerical instability. Hence, the same is true for the inhomogeneous system so that forward computation is not a meaningful procedure. To overcome this problem, further research is required.

Further research is also required for extending our results to block band transition matrices because there is no counterpart to Pringsheim's convergence criterion (Theorem 2) in the matrix-driven case.

Data Availability

No additional data were used to support this study.

Additional Points

Highlights. (i) Inherent instability of forward computation of absorption probabilities for certain transition structures is rigorously proven by establishing a structural result. (ii) Generalized-continued fractions are used to derive a decoupled difference equation. (iii) Based on this decoupled equation, a numerically stable algorithm is established.

Conflicts of Interest

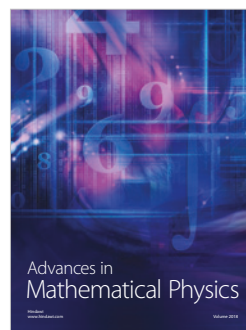
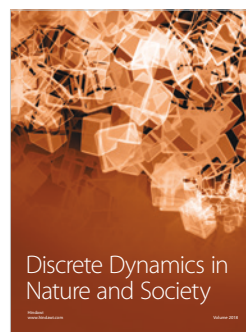
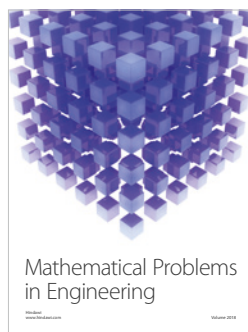
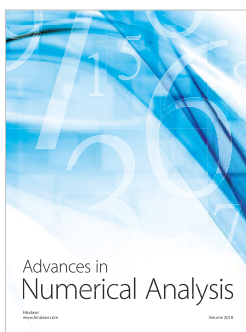
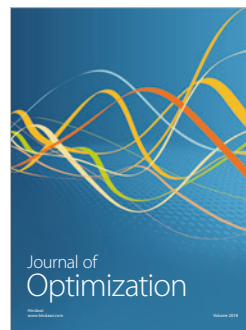
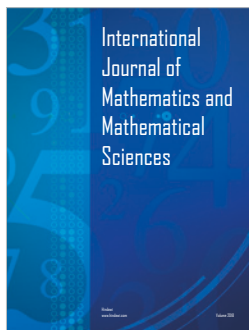
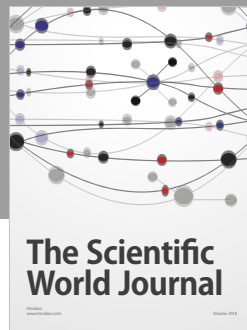
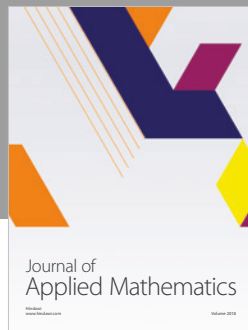
The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge support by the Open Access Publishing Fund of Clausthal University of Technology.

References

- [1] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, London, UK, 1993.
- [2] K. S. Miller, *Linear Difference Equations*, Benjamin, New York, NY, USA, 1968.
- [3] J. R. Cash, "A note on the numerical solution of linear recurrence relations," *Numerische Mathematik*, vol. 34, no. 4, pp. 371–386, 1980.
- [4] W. Gautschi, "Computational aspects of three-term recurrence relations," *SIAM Review*, vol. 9, no. 1, pp. 24–82, 1967.
- [5] W. Gautschi, "Zur Numerik rekurrenter Relationen," *Computing*, vol. 9, no. 2, pp. 107–126, 1972.
- [6] S. W. Lozier, "Numerical solution of linear difference equations," Report NBSIR 80–1976, National Bureau of Standards, US Dept. of Commerce, Washington DC, USA, 1980.
- [7] R. M. M. Mattheij, "Characterizations of dominant and dominated solutions of linear recursions," *Numerische Mathematik*, vol. 35, no. 4, pp. 421–442, 1980.
- [8] H. Baumann and T. Hanschke, "Inherent numerical instability in computing invariant measures of Markov chains," *Applied Mathematics*, vol. 8, no. 9, pp. 1367–1385, 2017.
- [9] T. Hanschke, "Markov chains and generalized continued fractions," *Journal of Applied Probability*, vol. 29, no. 4, pp. 838–849, 1992.
- [10] P. Van der Cruyssen, "Linear difference equations and generalized continued fractions," *Computing*, vol. 22, no. 3, pp. 269–278, 1979.
- [11] O. Perron, "Über lineare Differenzen- und Differentialgleichungen," *Mathematische Annalen*, vol. 66, no. 4, pp. 446–487, 1909.
- [12] T. Hanschke, "Über die Minimallösung der Poincaré-Perronschen Differenzengleichung," *Monatshefte für Mathematik*, vol. 112, no. 4, pp. 281–295, 1991.
- [13] T. Hanschke, "Ein verallgemeinerter Jacobi-Perron-Algorithmus zur Reduktion linearer Differenzengleichungssysteme," *Monatshefte für Mathematik*, vol. 126, no. 4, pp. 287–311, 1998.
- [14] P. Levrie and A. Bultheel, "Matrix continued fractions related to first-order linear recurrence systems," *Electronic Transactions on Numerical Analysis*, vol. 4, pp. 46–63, 1996.
- [15] O. Perron, *Die Lehre von den Kettenbrüchen II*, Teubner, Stuttgart, Germany, 1957.
- [16] P. Levrie, "Pringsheim's theorem for generalized continued fractions," *Journal of Computational and Applied Mathematics*, vol. 14, no. 3, pp. 439–445, 1986.
- [17] O. Perron, "Über die Konvergenz der Jacobi-Kettenalgorithmen mit komplexen Elementen," *Sitzungsber. Akad. München (math.-phys.)*, vol. 37, pp. 401–482, 1907.
- [18] S. Karlin and H. M. Taylor, *A Second Course in Stochastic Processes*, Academic Press, New York, NY, USA, 1981.



Submit your manuscripts at
www.hindawi.com